SciencesPo
LABORATOIRE INTERDISCIPLINAIRE
D'ÉVALUATION DES POLITIQUES PUBLIQUES

LIEPP Working Paper

April 2019, nº88

When in Rome…

on local norms and sentencing decisions

David ABRAMS
University of Pennsylvania
dabrams@law.upenn.edu

Roberto GALBIATI
CNRS-Sciences Po and CEPR
roberto.galbiati@sciencespo.fr

Emeric HENRY
Sciences Po and CEPR
emeric.henry@sciencespo.fr

Arnaud PHILIPPE
University of Bristol
arnaud.philippe@bristol.ac.uk

How to cite this publication :
David ABRAMS, Roberto GALBIATI, Emeric HENRY and Arnaud PHILIPPE, **When in Rome… on local norms and sentencing decisions**, *Sciences Po LIEPP Working Paper* n°88, 2019-04-17.

When in Rome...
on local norms and sentencing decisions*

David Abrams[1] Roberto Galbiati[2]

Emeric Henry[3] Arnaud Philippe[4]

April 17, 2019

Abstract

In this paper, we show that sentencing norms vary widely even across geographically close units. By examining North Carolina's unique judicial rotation system, we show that judges arriving in a new court gradually converge to local sentencing norms. We document factors that facilitate this convergence and show that sentencing norms are predicted by preferences of the local constituents. We build on these empirical results to analyze theoretically the delegation trade-off faced by a social planner: the judge can learn the local norm, but only at the cost of potential capture.

[1]University of Pennsylvania

[2]CNRS-SciencesPo and CEPR

[3]SciencesPo and CEPR

[4]University of Bristol

*We wish to thank Charles Angelucci, Alberto Bisin, Patrick Le Bihan, Olivier Marie, Aurelie Ouss, Salvatore Piccolo, Shanker Satyanath, Joel Van Der Weele, and seminar participants at the University of Amsterdam, University of Bristol, NYU, Oxford University, University of Pennsylvania, University of Bergamo, Collegio Carlo Alberto, University of Paris Ouest, Erasmus University, CELS and IAST-Toulouse for useful comments and discussions. Kriti Mahajan, Chloe Nibourel, Alyssa Staats and Anton Strezhnev provided excellent research assistance. Ryan Fackler is owed tremendous thanks for conveying his extensive knowledge of the data. Roberto Galbiati acknowleges the support of the LIEPP through a public grant overseen by the French National Research Agency (ANR) as part of the Investissements dAvenir program LIEPP (reference: ANR-11-LABX0091, ANR-11-IDEX-0005-02)

1 Introduction

The ancient Romans appreciated the merits of a flexible interpretation and application of the law, as evidenced by the Latin maxim *summum ius, summa iniuria* (Cicero, de officiis, 30), meaning "supreme law, supreme injustice." In modern legal practice, legislators still grapple with the necessity of granting judicial autonomy to allow for flexibility while setting forth legal rules. This autonomy means that sentences will vary along at least three dimensions: case, judge, and location. At the case level, judges take into account specific characteristics, such as mitigating or aggravating factors, in assigning criminal sentences. Sentences also vary by judge, due to differences in ideology, or potentially capture by local interests. Finally, sentences may vary spatially, due to local sentencing norms. While judges' biases have been extensively studied (Abrams et al. 2012, Lim et al. 2015, Cohen and Yang, 2018 and Schanzenbach and Tiller, 2007), we know little about the existence, magnitude, and determinants of spatial variation in sentencing.

This lack of evidence is primarily due to the empirical difficulty of disentangling geographical differences in sentencing norms from spatial variation in unobservable crime characteristics or spatial clustering in judges' preferences. Analyzing a unique setting where judges must rotate across judicial districts, we find that judges arriving at a new court gradually converge to the local sentencing averages – appearing to follow the well known proverb, "when in Rome, do as the Romans do." We thus both demonstrate the existence of local sentencing norms and document how judges dynamically adapt to them.

Later in the paper, we discuss the implications of these results for the question of optimal delegation of sentencing to judges. Judges, when granted autonomy, can obtain information about case specificities, but also, as we emphasize in our main contribution, about local norms, information that is costly for the planner to obtain. Letting judges adapt to case specific details is clearly socially desirable if the planner wants to establish a precise gradation of sentences. Adaptation to local norms of behavior is a more contentious issue. On the one hand, the planner has a set of guiding principles that should not be affected by local preferences. However, the planner may also want to at least partially adapt to local norms to limit the cost of enforcement, which can increase when sentences conflict with local preferences (Acemoglu and Jackson, 2017 and Hay and Shleifer, 1998). We present a theoretical model, building on the empirical evidence and including these trade-offs, that compares the primary instruments used in practice to constrain judges: sentencing guidelines, rotation, and judicial elections.

Our empirical analysis exploits the unique institutional features of North Carolina's Superior Court system and a rich dataset including all criminal justice decisions made in these courts between 1998 and 2010. In this system, judges are subject to elections as well as regular spatial rotation. The State is divided into eight divisions, each of which are subsequently divided into a variable number of districts. Judges are elected in one particular district, but they do not stay there permanently. Every six months, in January and July, they have to change district within their division, according to a schedule determined by the State Chief Justice.

The starting point of our analysis, and one of our main contributions, is to show that judges gradually adapt to local sentencing practices when they arrive in a new district. Specifically, for each judicial district, we compute the average sentence in a crime category given by "senior judges" (judges elected before 1998), who presumably already have good knowledge of the local conditions. We refer to this average sentence at the district crime level as the "local sentencing norm". Restricting the data to decisions made by "junior judges" (judges elected after 1998), for whom we observe the full history of sentencing, we show that the absolute value of the distance between the chosen sentence and the local sentencing norm decreases with the number of cases examined by the judge in the district. Our results imply that comparing the hundredth case to the first in any given district, the judge gives a sentence 24 days closer to the local sentencing norm.

By documenting the evolution of judicial sentencing as a function of time spent in a district, our methodological approach shows both the existence of a local sentencing normal and gradual adaptation to it. In particular, our identification strategy overcomes several challenges, that we discuss in detail in Section 4.1. A first concern is that judges recently arrived in a district might be assigned cases more distant from the local norm. We provide evidence that this is not the case by showing that observable characteristics, such as crime type or the defendant's sex or race, do not evolve over time spent in a district. Second, selection might explain our results if judges are able to affect the time they spend in the different districts. To eliminate this possibility, we focus on a balanced sample of judges, restricting our working sample to the first 400 decisions in a district for judges making at least 400 decisions. Finally, by using judicial rotation as our source of identification, we can separate adaptation to a local norm from learning about the law in general.

Further, we differentiate decisions made by judges working in their home district (which we call the home district) from those made outside and show that convergence only happens for those in the second category. This is consistent with the idea that judges are already aware of local sentencing practices in their home district. Digging into the details of the learning process we document that early sentences made in a district - before learning takes place - are partly determined by the sentencing norm that prevails in the judge's home district, an indicator of the judge's intrinsic preferences. When judges from relatively tougher (more lenient) home districts rotate to a new location, their sentencing converges to the new local sentencing norm from above (below).

Our analysis, exploiting assignment of judges to districts, also allows us to establish that spatial variation in sentencing is much larger than judge-level variations. Regressing sentences on judge and district fixed effects while controlling for crime characteristics, we find that the standard deviation in district fixed effects is twice as large as that of judge fixed effects. This suggests that, while the literature has mostly focused on judicial biases, a significant part of the variation attributed to judges could in fact be due to local characteristics.

After having shown the existence of a local norm of sentencing, in Section 4.3 we examine what factors seem to explain variations in the norm. There are three main candidates. Local norms could reflect constraints that district judges take into account (such as constraints on the police force or on prisons), preferences of the population (norms of behavior and customs) or just a court-specific culture unrelated to fundamentals. Our analysis suggests that the main driver is local preferences: controlling for district and crime fixed effects, variations in the prevalence of a particular crime in a district decreases average sentences for that crime in that district. We interpret this as the judge adapting to the norm of greater tolerance towards certain types of crimes. In a similar vein, we show a district's sentencing norms are correlated with a community's votes in referenda on criminal justice questions. Site-specific variables measuring resource constraints, in particular prison overcrowding, and those trying to capture culture, do not appear to play a role.

Our empirical findings show that the judge, when granted autonomy, adapts both to case characteristics (such as mitigating factors) as well as local preferences. While adapting to case characteristics is welfare-enhancing for a planner wanting to have a gradation of sentences, adaptation to local norms is a more contentious issue. We build a theoretical model to approach this problem and to characterize optimal delegation of sentencing to judges.

In the model, a representative citizen has a preferred sentence that varies by district. The planner has a statewide preference and would ideally not adapt to local preferences, but faces a cost of enforcement that depends on the distance between the sentence chosen and the preferred sentence at the district level.[1] Given this constraint, the planner prefers to move in the direction of the local norm, but only partially. However, the planner observes neither case-specific characteristics nor the local norm, and thus delegates the decision to a judge, who, for each case, can observe the specific details, and can also observe local preferences

[1] Such a constraint is also present in papers discussing the interaction between law enforcement and social norms and customs (Acemoglu and Jackson, 2017 and Hay and Shleifer, 1998), noting that when the two are too distant enforcement costs increase.

for a cost. The judge also wants to adapt to the local norm, because, for instance, she risks having her decision appealed if she does not. In terms of sentencing preferences, the judge can either be aligned with the planner with some probability, or prefer tougher sentences.

Within this environment, we build on our empirical findings by comparing social welfare under three main instruments of judicial oversight: judicial rotation, sentencing guidelines, and judicial elections. In the model, the planner's preferred policy partially adapts to the local norm, taking enforcement costs into account. The judge, under rotation, also partially adapts, as observed in the data. Our main finding is that whenever the uncertainty on unobservable crime characteristics is large, rotation will be the socially preferred instrument. Sentencing guidelines in this case naturally perform poorly since they cannot condition on ex ante unobservable characteristics. Elections also perform poorly, because, at equilibrium, all types of judges pool on the same sentence to avoid revealing their type, and accordingly ignore information on case specific conditions. However, when the uncertainty is small, sentencing guidelines may be preferred, as they avoid the cost of having a biased judge decide.

The trade-off between information gathering and excessive adaptation to local interests is relevant in many contexts beyond the judiciary. In the corporate context, such as consulting, local representatives are often used to gather valuable local information and forge relationships with particular clients. However increased familiarity can also lead to favoring client interests over those of the company if relationships grow too close. To counteract this tendency, employees in this context are often rotated.

Similarly, over time, local politicians acquire intricate knowledge of the problems facing their jurisdiction but are also susceptible to capture[2] by local power brokers. Elections and term limits are two approaches to address these problems, with known limitations. Our investigation of this phenomenon in the judiciary is both interesting in its own right, and also allows us to empirically describe the general trade off. We exploit the fact that we observe the full set of criminal decisions made by a judge over time. In general, the full work output of many other professionals is not readily observable nor quantifiable.

The paper is related to different strands of literature. First, it relates to studies on judicial decision-making. As mentioned above, most of the recent research in this area aims at explaining inter-judges variation in criminal justice sentencing. Some papers have shown how judges' personal characteristics help explain their decisions. For instance, Lim et al. (2016) study the influence of judges' political orientation and demographic characteristics on criminal sentencing decisions in Texas. The authors document substantial heterogeneity

[2] Throughout this paper we use capture to mean influence which may cause judicial sentences to deviate from the social optimum, but are not necessarily illegal.

in sentencing harshness across judges. Cohen and Yang (2018) show that judge's political orientation contributes to gender and racial disparities in the US federal courts. Yang (2014) shows that, when they are not constrained by sentencing guidelines, judges discriminate more against minorities in criminal justice decisions. Abrams, Bertrand and Mullainathan (2012) find evidence that heterogeneity in judicial decisions is driven by defendant race. Berdejo and Chen (2017) show how judges are influenced by the political climate modifying their decisions when presidential elections are closer. There is also a substantial theoretical literature that studies how behavioral biases and individual motivations shape judicial decisions in both tort and criminal law (Gennaioli and Shleifer, 2008 and Bordalo et al. 2015).

The literature on spatial variations is much more limited. Yang (2014) documents spatial variation in sentencing across the US but cannot distinguish between crime characteristics and different local practices. Lim et al. (2016), using data from Texas state district courts where judges overlap in different districts, find results in contrast to ours: judge fixed effect tend to be more variable than district fixed effects, and judges do not seem to adapt to behavior of other judges in their district. One important distinction with their work is that, unlike in our institutional setting, judges sort in particular districts through the electoral process. Ichino et al. (2003) show that Italian labor judges, take into account local constraints and adapt their reintegration decisions to labor market circumstances. One of our main contributions is also to describe how judges learn about local sentencing norms. To the best of our knowledge, the economic literature empirically documenting learning in judicial decision making is scarce, but Coviello et al.(2018) are a notable exception documenting how judges learn by doing how to treat similar cases.

Beyond judicial decision making, our paper contributes to a recent literature explaining spatial variations in the provision of public services. In particular, our approach is similar to two recent papers explaining spatial variation in medical doctors' diagnostic practices in the US (Molitor, 2018 and Finkelstein et al. 2016), although in our setting, the usual challenges to identification due to endogenous sorting are overcome by the institutional design.

Finally, our study contributes to the literature studying the effects of policies aimed at reducing the capture of public officials. The closest examples are recent papers analyzing the effects of term limits on politicians behavior (Coviello and Gagliarducci, 2017 and Dal Bo and Rossi, 2011) and how UK tenured judges respond to political pressure (Blanes i Vidal and Leaver, 2011).

The paper is organized as follows. In Section 2, we describe the institutional setting. In Section 3, we present the data and show motivating evidence that spatial variations are larger than judge-level variation. Section 4 discusses our identification strategy, presents the evidence on adaptation of judges to local practices, and describes factors that correlate with these local sentencing norms. Finally, in Section 5, we use the empirical findings to build a theoretical model that characterizes optimal delegation of sentencing to the judge.

2 Institutional Setting

In North Carolina, the Superior Courts have general trial jurisdiction over civil and criminal matters. The Superior Court system uniquely combines judicial elections, judicial rotation, and sentencing guidelines. While sentencing guidelines and judicial elections are quite common, since the majority of States use elections for at least some judicial positions, rotation is a relatively unique feature of modern state courts. It was, however, more common in the past, when judges would "ride circuit" to provide justice to more rural areas.[3]

In practice, the State is divided into eight divisions, each of which are partitioned into districts (see figure 1 in the Appendix A).[4] Elections take place separately in each district, and rotation is then organized at the division level, as we describe below.

Elections

The selection of Superior Court Judges takes place via non-partisan elections. Judges are elected in a home district for an eight-year term. If a vacancy arises in the middle of a term, the governor fills the vacancy by appointment, which is effective until the next general election. There are no term limits; after their term ends, judges can run again for elections. Although they are required to run again in the same district, they usually end up doing so.[5]

Rotation

Superior Court judges elected in each district have the obligation to rotate every six months, in January and July, within the division where their home district is located. This rotation rule was established in the North Carolina Constitution of 1868. The primary motivation for the rotation system is to avoid capture by the local community. As expressed by a Superior Court Judge in the local press,"because judges are elected, building alliances through campaigns and asking for campaign contributions largely with people in their local districts, rotation reduces bias and the perception that contributors have better access or influence."[6] A similar motivation is presented in Bobbitt (1948) who also emphasizes the

[3]South Carolina and Nevada are the only two other states that still use some sort of judicial rotation.

[4]Before 1999, the State was divided into 4 divisions, which was then increased to 8 to reduce the distance judges had to travel.

[5]In our sample, only 2 judges choose to run in a different district.

[6]See " "Riding the circuit: Traveling judges program ensures impartiality, comes with cost", in the Times News. He adds, "Rotation mitigates the building of inappropriate or overly-familiar relationships and enhances the perception that a truly impartial judge will be presiding. In this way, judicial independence is enhanced, and an independent judiciary is a principle of the highest value."

benefit of avoiding an excessive connection between judges and particular lawyers.[7] The additional argument provided in Bobbitt (1948) is that, if the judge is too immersed in the local community, he might act "on the basis of fixed ideas of his own or on the basis of local reports rather than on the basis of the evidence before him."

The Chief Justice of the North Carolina Supreme Court establishes the rotation schedule. There are no formal rules that the Chief Justice needs to follow, even though there is a common understanding that judges should return to their home district (the district of their election) at least once every two years. Judges can present motivated objections to their particular assignment. This official rotation policy is in fact respected in practice. Using only observations in our data, we plot in Figure 2 the average probability of moving district by month, where the district of work in a month is defined as the district where a majority of cases are decided. Consistent with the rotation policy, a large majority of judges move to a new district in January and July.[8]

Sentencing guidelines

Judges' decisions for criminal cases in North Carolina are structured by the State's sentencing laws that went into effect in 1994.[9] To choose a sentence for a felony case, the judge first has to determine the offense class to which the felony belongs. She then establishes the prior criminal record of the offender. There are ten offense classes ranked by severity and six different criminal history levels. Once the offense class and the criminal history of the offender are determined, the judge must establish which aggravating and mitigating factors apply to the case. Given these choices, the case falls into a cell in the sentencing grid. Each cell defines the range of possible minimum sentences. The judge has to select an appropriate sentence within this minimum sentence range (see Figure 10 in Appendix D). For instance, an offender convicted for an armed robbery (a class D offense) and having prior record of two offenses, faces a default sentencing range of minimal sentences (called presumptive sentence range) ranging from 79 to 97 months. If the judge finds aggravating circumstances, the range shifts to 97-121, whereas mitigating circumstances move it to 58-78 months. Thus, depending on the circumstances, and the criminal history, the minimal sentence for an armed

[7]"If it be true, or if it be thought by the litigants or by the public, that the judge, whether through personal friendship with lawyer or with litigant, or through previous experience with lawyer or with litigant, or through political, business, social or other connections, will be swayed, consciously or unconsciously, by considerations other than the law and the evidence in the particular case, the prestige and usefulness of the judge is greatly impaired. It may be well to recall that, in every controversy before the court, each decision a judge makes is a decision against some person or persons."

[8]The frequency of moves in these two months is larger than 80 percent. There are some deviations, especially in the one month preceding and following the scheduled moves, presumably to smooth transitions.

[9]The sentencing procedure is detailed in the North Carolina 'Structure Sentencing Training and Reference Manual' available at https://www.nccourts.gov/assets/documents/publications/sstrainingmanual09.pdf

robbery can actually vary between 58 and 121 months.

Finally, the law may stipulate a sentence disposition for each combination of offense class and prior record level: the sentence may be active (i.e. the offender is under custody) or suspended. If a defendant receives an active sentence, he must serve the entire minimum sentence defined above. For certain class-record combinations, the judge has to decide whether the sentence should be active or suspended.[10]

3 Data and descriptive statistics

3.1 Data

Sentencing data

Our data comes from the North Carolina Administrative Office of the Courts and includes the universe of felony cases decided in North Carolina superior courts from 1998 to 2010. For each decision, we know the week of the decision and the identities of the defendant, defense attorney, district attorney, and judge. The data also includes the main demographic characteristics of the defendant as well as their criminal charges. The dataset includes 343,776 sentencing decisions with final disposition dates between 1998 and 2010. It is worth noting that our main unit of analysis (a case) is defined by aggregating all outstanding charges for a defendant that are disposed of at one time. The construction of the data is described in more detail in Appendix D.

In the main analysis, our main sentencing variable is the minimum active sentence chosen by the judge as in the procedure described in Section 2. If multiple cases are disposed of at the same time for the same defendant, we consider the maximum of the minimum sentences, i.e the decision made for the most severe offense. In Appendix B, we consider the robustness when using as dependent variable the total sentence, both active and non.

Master schedule and electoral data

We also obtained the master schedule produced by the Chief Justice that describes the assignment of each judge across the districts. We use this data to guarantee that we correctly identify the judge; we restrict the data to observations where the judge is in the division recorded in the master schedule. The identity of the judge is then used to match with the result of judicial elections.

[10]As in the rest of the U.S., the vast majority of criminal cases in North Carolina are resolved via plea bargain, where the sentence is agreed to between the prosecution and defense. Even though the judge is often not directly involved in plea bargains, she is required to approve the sentence, and thus still exerts a great deal of influence over the outcome (Abrams and Fackler, 2018 and Mnookin and Kornhauser, 1979).

Other sources

We obtained voting data from North Carolina's Board of Elections. First, results on voting by district for all presidential elections held during our time period. Second, data on voting in four referenda held before, during and after our time period, that involved proposals related to the judicial system (for details on these proposals, see Appendix D). To compute jail overcrowding, we used the Annual Survey of Jails (ASJ) Data Series, which is collected by the Bureau of Justice Statistics.

3.2 Descriptive statistics

Table 1 reports descriptive statistics. In panel A we present case and judicial characteristics, separately for 70 junior judges (those elected or appointed after 1998) and 91 senior judges (those elected or appointed before 1998). This distinction will prove essential in our empirical analysis.

More than 80 percent of defendants are young males, (the mean age is 31 years), and more than half of them have a criminal record. Half of defendants are black, and races other than black and white represent less than 10 percent of cases. Judges spend around half of their time in their home district, senior judges have average tenure in office of 14 years, and junior judges have an average tenure of 7 years. The average sentence is about 500 days, while the active sentence is about one year.

In Panel B, we present average characteristics of the 50 districts in North Carolina. There are, on average, 11 judges and 23 district attorneys who take at least one decision in the district. Naturally, the pool of defense attorneys is larger, with an average of 61 per district. The rest of panel B presents summary statistics of political preferences and population characteristics that we'll use in our analysis of the determinants of the sentencing norms. North Carolina tends to vote more Republican.

3.3 Spatial variation in sentencing decisions

The premise of our study is that spatial variation in sentencing exists. We document in Figure 3 that this is indeed the case. Sentences vary widely not only by district but also by crime type. For instance, sentences for violent crimes range from an average of around 400 days in lenient districts to approximately 800 for the harshest.[11]

[11]Forsyth County, which contains Winston-Salem, is the outlier in all panels of Figure 3. In private conversation, a local official suggested that local prosecutors have a norm of aggressive prosecution, especially among most serious crimes.

Several factors could cause this spatial variation. First, the distribution of crimes committed or the unobserved crime characteristics within a crime type could differ across districts. Second, the pool of judges differs by division and could vary in harshness.[12] We can account for these variations using the following specification:

$$S_{ijdc} = \beta X_i + \delta_j + \gamma_d + \nu_c + \epsilon \tag{1}$$

where S_{ijdc} is the sentence decided by a judge j, in district d for given crime c and i a case indicator, δ_j, γ_d and ν_c, respectively judge, district and crime fixed effects and X_i are case characteristics, such as age, gender and race of the defendant.

In the left panel of Figure 4, we present the distribution of district fixed effect γ_d where equation (1) is estimated using only the decisions made by senior judges, who have acquired good knowledge of each district.[13] After controlling for case characteristics and judge fixed effects, large variations across districts remain. Excluding two outliers, fixed effects by districts still vary within a 200 days band.

These results also allow us to examine how much of the deviation in sentencing from a crime level average can be explained by judicial versus geographic variation. As highlighted in the introduction, distinguishing spatial and judicial fixed effects relies on our institutional setting, where judges rotate across districts. In the right panel of Figure 4, we present the judicial fixed effects for the senior judges who take more than 500 decisions and visit at least two districts during our observation period. Variability in district fixed effects is approximately twice as large as variability in judicial fixed effects. This finding suggests that spatial variation in sentencing is indeed an important feature per se of the criminal justice system over and beyond the variation explained by judges' fixed effects. Understanding spatial variation is thus crucial for describing the system's design.

Even though in estimating equation (1), we control for crime fixed effects (with 615 categories) and for observable characteristics of the defendant, the possibility remains that variations across districts is explained by systematic geographical differences in crime characteristics, unobserved by the econometrician but observable to the judge.[14] The local fixed effect could therefore include two components: first, differences in average unobservable crime characteristics, second, different behavior across space in sentencing for the same crime characteristics, which we call the local sentencing norm. Our identification strategy, presented

[12]Even though we have rotation and thus constraints on sorting by judges, judges can still sort across divisions and might be able to affect the time they spend in each district.

[13]The separate use of senior judges will be described in more detail when we discuss identification in Section 4.

[14]This explanation seems unlikely, since observable characteristics such as race, gender and sex do not reduce the geographical variance much.

in the next section, allows us to both show the existence of this sentencing norm and the gradual adaptation of judges to it.

4 Adaptation to local sentencing norms

4.1 Identification

Two features of the institutional setting in North Carolina provide important benefits for identification. First, we have access to a relatively large dataset with a significant number of judges arriving in new jurisdictions, at different points in their career, all under the same State law. Second, we can distinguish decisions made by judges in their home district versus other decisions made elsewhere, which helps us to distinguish adaptation to a norm from learning about the law.

We divide our sample of judges in two groups: junior and senior judges. For junior judges, the entire sentencing history is observed, including the number of decisions made in each district. We use this set of judges to construct our working sample. We use the senior judges to compute the local effect $LocSent_{dc}$, defined as the average sentence given by senior judges in district d for a given crime category c.[15] The idea is that, if the local effect includes a local sentencing norm component, these judges will have already spent more time in the district, allowing them to have better knowledge of this norm. We show in Section 4.2 that our main results are robust to changes in the definition of $LocSent_{dc}$, in particular using only senior judges elected in the district to define the local effect.[16]

In our main specification, we examine how sentences evolve compared to $LocSent_{dc}$:

$$\begin{aligned} |S_{ijdc} - LocSent_{dc}| &= \alpha_1 Order_{jd} + \alpha_2 Order_{jd} * 1_{ElecDistr} \\ &+ \alpha_3 1_{ElecDistr} + \alpha_4 Order_j + \beta X_i + \epsilon \end{aligned} \quad (2)$$

where S_{ijdc} is the sentence decided by a judge j, in district d for given crime c and i a case indicator. $Order_j$ counts the number of cases treated by judge j up to date t (judge experience) and $Order_{jd}$ counts the number of cases treated by judge j up to date t in the district d (judge local experience).[17] Finally $1_{ElecDistr}$ measures whether the judge is

[15] An alternative would be to extract this variable as a crime * district fixed effect in a modified version of equation (1). However, this would imply extracting 615*50 fixed effects and most of them would be very imprecisely estimated.

[16] If the sample of judges used to define $LocSent_{dc}$ have not fully learned the norm, this would create a noisy measure of the local effect. It is reasonable to think that elected judges would have better knowledge of the local effect, hence the robustness check.

[17] We will also run quantile regressions of equation (2) where the dependent variable is taken without

sentencing in her home district.

The main parameter of interest is α_1, which measures the impact of the number of decisions made in a district on the distance to the local effect, and provides two essential pieces of information. First, under the assumption that the unobservable characteristics of the case (to the econometrician) are always observed by the judge, a value of α_1 significantly different from zero (i.e an evolution in time of the judge's sentencing behavior) indicates the existence of a local sentencing norm. Indeed, if variations in sentencing norms were only driven by variations in unobservable case characteristics, since the judge observes them, there would be no evolution in time and α_1 would be zero. A negative value of α_1 indicates gradual adaptation to the local sentencing norm as a function of number of cases seen in a district.

The first potential identification challenge is related to assignment of cases to judges. If the judge is assigned cases more distant from the average case when she first arrives in a district, α_1 would potentially capture this assignment dimension rather than adaptation to the local sentencing norm. One way to address this concern is to examine how observable characteristics vary as a function of the order of decisions in a district. We examine this in Table 2 where we estimate equation (2) using different case characteristics as the dependent variables. In panel A, we report the coefficients when we use the defendant characteristics as dependent variables, and in panel B, the crime categories. Reassuringly, we find that the order of the sentences made by a judge in a given district does not have a significant impact on the type of case or the characteristics of the defendant.

The second potential challenge is to differentiate between adapting to the local sentencing norm and merely learning about the law. This concern is alleviated by the fact that different judges arrive in the same district at different points in their careers, due to the rotation system. In the specification, we separately include judicial experience in the district and overall experience of the judge, with the latter effect being measured by parameter α_4.

We also allow for the fact that judges likely already know the norms in their home district by interacting the variable $Order_{jd}$ with a home district dummy. Under the reasonable assumption that a judge does not learn more quickly about the law in her home district compared to outside, if we find α_2 to be significantly different from zero, this would suggest that we are indeed measuring adaptation. In particular, we would expect to find α_2 positive, and roughly of the same absolute magnitude as α_1, consistent with the idea that the judge already knows the local sentencing norm in her home district.

The third potential challenge is the risk that judges can affect the time they spend in

the absolute value. If judges come closer to the local effect over time, we expect the coefficients for lower quantiles (resp. higher) to be positive (resp. negative).

a district. As explained in Section 2, the master schedule is controlled by the State Chief Justice. While we have no evidence of favoritism in assignment, judges are allowed to make scheduling requests. If endogenous sorting does occur, finding α_1 to be significant (i.e. the fact a judge behaves differently for early decisions versus late) could capture the fact she is able to spend more time in districts where she is closer to the norm. To address this concern, we use a fully balanced dataset. We homogenize the dataset by focusing on the first 400 decisions made by a judge in a district where the judge takes at least 400 decisions.[18] This restriction leaves us with a dataset containing 48 different judges.

4.2 Results: adaptation to the local sentencing norm

Before presenting the results of the estimation of equation (2), we first provide graphical evidence of adaptation over time to the local sentencing norm. Figure 5 presents the results of an OLS regression of the following model, separately for different quantiles:

$$S_{ijdc} - LocSent_{dc} = \sum_{k=1}^{7} \alpha_k * \mathit{1}_{order=k} * \mathit{1}_{ElecDistr} + \sum_{k=1}^{7} \beta_k * \mathit{1}_{order=k} * (1 - \mathit{1}_{ElecDistr}).$$

The figure shows how sentences evolve with the number of decisions made in a district. We separately present the results for decisions made in the judge's home district (right panel) and in other districts (left panel). There is no clear evolution of sentence length for judges hearing cases in their home district (right panel). This contrasts markedly with what happens when a judge is outside her home district (left panel). In this case, the left panel of Figure 5 shows that the gap between the 90th and the 10th percentile of the distribution of distance between the sentence and the local target ($S_{ijdtc} - LocSent_{dc}$) reduces over time,[19] while the median of the distribution remains very stable.[20]

These initial graphical results not only show the existence of and adaptation to a local sentencing norm, but the contrast between the left and right panels also validates our identification strategy. Those in their home district a priori know the local target better as they already had to campaign in the district, and presumably live there. The fact that they do not change their sentencing behavior over time confirms that we observe adaptation to a local norm rather than learning about the law, which should occur uniformly across all districts that are subject to the same legal rules.

[18] We also perform robustness checks in Section 4.2, where we use thresholds of 500 and 300 decisions.

[19] Note that the same judge can take decisions in different parts of the distribution.

[20] In Figure 8 in Appendix B we reports the same results described above when using as a dependent variable the total sentence rather than only the active sentence. The patterns described above remain substantially unchanged.

The graphical results are confirmed in Table 3 where we estimate variations of equation (2) using both OLS and quantile regressions, with standard errors clustered at the judge level. Column (1) estimates the equation, removing the absolute value from the dependent variable. For this specification we find no effect of the number of cases in a district on the distance between the sentence and the local norm. This can be seen as indirect evidence of the absence of strategic allocation of cases over time. If more (less) severe cases with respect to the local sentencing norm were systematically assigned to judges with less experience in a district we would expect a positive (negative) and significant effect of the order of decisions in the district. Column (7) estimates equation (2), where the dependent variable is the absolute value of the distance between the sentence and the norm, and shows that α_1 is indeed negative and significantly different from zero. The magnitude of the effect is large: the value of α_1 implies that, comparing a judge's 100th case in a district to the first, the sentence will be 24 days closer to the local norm, when she is outside her home district. Parameter α_2 is positive and slightly smaller in absolute value than α_1. The sum of the two parameters is not statistically different from zero, confirming that this adaptation is absent for decisions made in the judge's home district.

Columns (2) through (6) present the results of quantile regressions with the difference between the sentence and local norm as the dependent variable. The overall results point to convergence to the local norm throughout the judicial sentencing distribution. The coefficient of interest, α_1, is positive for the reported percentiles below 50 and negative above. For each quantile, α_1 and α_2 are of opposite signs. Magnitudes are particularly large for the extreme quantiles: for the 90th percentile, comparing a judge's 100th case in a district to the first, the sentence is 66 days closer to the local norm. Figure 6 presents a telling graphical representation of these results, where we plot the coefficients for different quantiles of the distribution. The coefficient is positive and significant for the lower quantiles and negative and significant for the higher quantiles.[21]

We consider a number of robustness checks of our main results in Table 4. First, we examine whether our results are sensitive to the definition of the local sentencing norm. In the main specification the variable $LocSent_{dc}$ was computed using only decisions made by senior judges, with the idea that they already had time to adapt to the norm.[22] We show in column (1) that estimating equation (2) when using an even more conservative definition of $LocSent_{dc}$, restricting ourselves to senior judges in their home district, yields comparable estimates of α_1 and α_2. In column (2), we use a more liberal definition of the norm, that

[21]Figure 9 in Appendix B shows equivalent results when including suspended sentences.

[22]The sample size drops with this specification because for some district*crime categories we do not have any decisions made by senior elected judges.

incorporates all decisions by junior judges after their 400th in a district (these decisions are used to compute $LocSent_{dc}$, but not to estimate the model). We find that the results still hold with this alternative definition of the local norm.

In order to rule out concerns of potential correlation between time spent in a district and judicial preferences, in our main specification we use a balanced sample of decisions, including just the first 400 decisions made by a judge in a district where the judge makes at least 400 decisions total. In column (3) we conduct the same exercise using the first 300 decisions and the first 500 in column (4). In both cases, the magnitude of the effects are very similar to those reported in Table 3 column (6). In column (5), we estimate (2) adding defendant races, sex, and age; column (6) includes judge fixed effects. The results are robust to the addition of these extra controls. This comes as no surprise since by focusing on a balanced sample, we rule out composition effects linked to the order of decision. Finally, in column (7) we use an alternative definition of sentence that adds suspended to active sentences. Again, the results hold substantially unchanged.

What determines the starting point of the learning process when a judge first arrives in a district? Figure 7 shows that when the sentencing norm in the judge's home district is higher (lower) than in the current district, the difference between the judge's sentences and the local average tends to be positive (negative) at the beginning of her tenure in the district. This effect disappears once the judge has examined a sufficient number of cases in the district. The key role of the home district could just be a proxy for the judge's preferences, i.e. she gets elected in a district with similar sentencing preferences.

Our main results show the existence of a local norm and gradual adaptation of judges to it. Our preferred interpretation of the adaptation process is that the judges want to adapt to the local norm, are initially unaware of it and gradually learn about it, case after case. We cannot rule out, however, that this process corresponds more to judges initially resisting the local norm and gradually giving in, for instance because local actors (e.g. attorneys, prosecutors) learn how to maneuver the judge. For exposition purposes, we will refer to the adaptation process as learning about the norm, but return to this point when discussing welfare in Section 5.

4.3 Local sentencing norm

The previous sections established the existence of a local sentencing practice and gradual adaptation of judges to this sentencing norm. Although we cannot establish the determinants of the sentencing norm in a causal manner, in this section we explore correlates of variables meant to capture the three main possible explanations for this norm. As mentioned in the introduction, the sentencing practices could reflect: local norms of behavior, local district resource constraints (on police or prisons) or, finally, a court-specific culture.

We measure the first dimension, local norms of behavior, in two ways. The idea of these local norms is that some illegal behavior might be more or less socially acceptable to the community.[23] Our first approach is thus to construct the district level prevalence of certain crimes, measured as the proportion of crimes in that district committed in a particular crime category.[24] Our second approach to measure local norms of behavior is to measure local political preferences. We use the results of the three presidential elections run during our time period (2000, 2004, 2008) as well as local referenda on judicial questions, described in Appendix D.

To measure the second possible determinant of the sentencing norm, we use one particular constraint, prison overcrowding. However, overcrowding is not a major issue in North Carolina, where only 20 of the districts have an issue of overcrowding. Even for these districts, the ratio of convicts to beds in close to 1, as described in Table 1. We might therefore be underestimating the impact of local constraints.

Measuring local culture at the courtroom level is more challenging since it is, by definition, the aggregation of characteristics of the permanent staff (attorneys and clerks) working in the district. Rather than attempting to measure it directly, we examine whether the concentration in the attorney or the district attorney market has an influence on average sentences. Specifically, we construct an Herfindhal Index of attorney or district attorney concentration at the district level. The index for attorneys is defined as

$$HI_d^a = \sum_{a \in \mathcal{A}_d} (N_{ad}/N_d)^2$$

where $\mathcal{A}_d$ is the set of attorneys working predominantly in district d, N_{ad} is the number of cases in d where a was the attorney and N_d is the total number of cases in district d. This index is calculated by year and averaged over the years in our sample. The Herfindhal Index for district attorneys is calculated in a similar fashion. If the local sentencing norm is defined by attorneys, a higher Herfindhal Index for attorneys should lead to lower sentences, since attorneys should presumably try to obtain lower sentences for their clients.

Results are presented in Table 5 where we regress the local sentencing norm $LocSent_{dc}$, on the determinants described above, controlling throughout for crime fixed effects. The first result is that an increase in crime c's prevalence in district d has a significant negative

[23]Research in criminology has presented evidence in line with this hypothesis (Eisenstein et al. 1988, and Ulmer and Johnson 2004), the idea is also supported by various discussions we had with defense attorneys and prosecutors.

[24]Specifically the variable is constructed as $100 * (cases_{dc}/cases_d) - 100 * (cases_c/cases)$.

effect on the sentencing norm, both controlling for district level controls (columns (3) to (5)) or including district fixed effects (column (2)). A 1 percent increase in the crime prevalence decreases the average sentence in the district by 4 days for that crime category. Note that the alternative interpretation is that the causality goes in the opposite direction: lower sentences for certain crimes in a district induce rational criminals to adapt to it. This explanation appears unlikely, since it would require very detailed knowledge of variation in the way judges behave locally.[25]

In contrast, neither local courtroom culture, captured by the concentration measures (introduced in column (3)), nor local constraints (introduced in column (4)) have an impact. In column (5) we introduce the political variables. The overall political leaning does not play a role. There seems to be a systematic link with the results of the 2004 and 2014 referenda. Note that the link is stronger when we consider in Table 6 in Appendix B, the sentencing norm calculated using total sentences rather than active sentences. Interestingly, the referendum of 2004 proposed to make judges accountable by fixing their first term at 2 years, while the 2014 referendum transferred power to the judge, by allowing criminal defendants to be sentenced by a judge rather than by a jury. The results in column (5) suggest that in districts where the judges are tougher, the citizens are more willing to transfer prerogatives to them. This is consistent with the idea that on average, citizen desire higher sentences than judges, which will relate to the next section's welfare discussion.

5 Judicial delegation

We have presented in the previous sections empirical evidence on several consequences of delegating the sentencing decision to the judge. First, as already shown in the literature, judges adapt to characteristics of the crime and criminal. Second, as shown in the main contribution of this paper, delegation allows judges to learn and adapt to the local sentencing norms. The results presented in Section 4.3 provide suggestive evidence that this sentencing norm is linked to local preferences.

This raises the question of why a social planner would want to constrain the freedom of the judge, since setting sentencing guidelines or rotation policies would decrease the speed of learning and the ability to adapt to local conditions. The answer to this question critically depends on whether adaptation to local norms is socially desirable. It is natural to think that the planner wants to stick to legal principles that are not abandoned in the face of local customs. Moreover, with delegation, the judge might end up favoring certain groups,

[25]Papers that find incentive effects of sentences (Drago et al. 2009 for instance), use large and salient legal changes in expected sentences, while our variations are small and not codified.

consciously or unconsciously, when spending too much time in a district, a concern at the root of the rotation policy.[26] However, adaptation can also be socially desirable if it allows easier enforcement of the law in an environment of higher social acceptance (as in Acemoglu and Jackson 2017).

This discussion highlights the trade-offs involved in delegating the decision to the judge and offers some rationale for planners constraining judicial autonomy with the use of tools such as sentencing guidelines, rotation, or elections. Building on our empirical findings, we thus consider the question of the optimal constraint on delegation in a model that captures the aforementioned main ideas.

5.1 A model of judicial delegation

Each period, a judge examines a case, indexed by time t. All players discount the future at the same rate δ. Each case has its own characteristics (of the crime and the criminal as shown in the data), summarized by the case severity denoted s_t distributed according to distribution f with mean 0.

The citizens

There is a representative citizen in each district d. The citizen has the following period utility when the sentence is set at S_t for case t:

$$u_c^d = -(S_t - (\theta + \beta_d + s_t))^2$$

where β_d is the local norm in district d, i.i.d drawn from distribution g with mean 0; θ is the target sentence in the state for a case of average severity $s_t = 0$. Thus, the desired sentence of the citizen in district d, for a case of severity s_t, is $S_c^d = \theta + \beta_d + s_t$. This representation fits the "penalty should fit the crime" heuristic. Aggravating conditions (higher s_t) and sentencing in tougher districts (higher β_d) lead citizens to demand higher sentences.

The planner

We consider a planner whose utility from sentencing is given by $-(S_t - (\theta + s_t))^2$, i.e who, for a given s_t, wants sentence $\theta + s_t$. In other words, the planner would like to stick to the basic principles of the law, embedded in θ, and not bend to local preferences. However,

[26] This is expressed in Bobbitt (1948): "If it be true, or if it be thought by the litigants or by the public, that the judge, whether through personal friendship with lawyer or with litigant, or through previous experience with lawyer or with litigant, or through political, business, social or other connections, will be swayed, consciously or unconsciously, by considerations other than the law and the evidence in the particular case, the prestige and usefulness of the judge is greatly impaired."

as in Acemoglu and Jackson (2017), when sentences are not aligned with local norms, this creates an additional cost of enforcement that the planner has to take into account. We introduce this idea of difficulties in enforcement in a reduced form way, assuming that there is a cost $\alpha(S_t - S_c^d)^2$ where S_c^d, defined above, is the desired sentence of citizens in district d. Parameter α measures the strength of this enforcement motive. Thus the utility of the planner is:

$$u_p^d = -(S_t - (\theta + s_t))^2 - \alpha(S_t - S_c^d)^2$$

Given these preferences, if the planner observed s_t and β_d, she would optimally choose sentence

$$S_p^d = \theta + s_t + \frac{\alpha}{1+\alpha}\beta_d$$

when $\alpha = 0$, the planner's desired sentence is $\theta + s_t$ and naturally, as α becomes large, the desired sentence converges to the representative citizen's preferred sentence $S_c^d = \theta + \beta_d + s_t$.

The judge

The judge, when the sentence deviates from the local norm, has to pay a cost $\gamma(S_t - S_c^d)^2$, corresponding either to the cost of disagreement with the local lawyers or the increased risk of having the decision appealed. It could also correspond to the risk of capture, which initially motivated the rotation policy, which would imply that spending time in a district and developing friendship and connections renders deviations from desired sentences more costly.

In terms of absolute preferences for sentencing, we assume the judge could be of two types. The regular type (with prevalence p) has the same preferences as the planner. The tough type (prevalence 1 - p) is biased in favor of longer sentences, with optimal sentence net of argumentation cost $\theta + s_t + \zeta$.

The per period utility of a judge of type j, who also receives a wage w per case independently of the sentence chosen, is thus given by:

$$u_j^d = -(S_t - (\theta + s_t + \zeta_j))^2 - \gamma(S_t - S_c^d)^2 + w$$

with $\zeta_j = 0$ for the normal judge and $\zeta_j = \zeta$ for the tough judge.

We assume that $\gamma > \alpha$; in other words, the judge's cost of deviating from the norm is higher than the planner's concern for enforcement costs.

Information

For each case t, the judge observes the severity s_t at no cost simply by listening to the case specifics. Consistent with the empirical evidence, we assume that if the judge is elected in a district she observes β_d at no cost. However, if she is not in her home district, the judge needs to pay a fixed cost C to learn β_d.

The planner, however, is unable to observe either β_d or s_t, and needs to delegate the decision to the judge. She has three instruments at her disposal to discipline the judge:

- **Sentencing guidelines**: fixed sentence S_g independent of the district and the unobservable crime characteristics.

- **Rotation**: the judge is required to move district every N cases.

- **Elections:** at the end of each period, there is a probability e that an election is run between the incumbent judge and a challenger. When an election is run, the citizen can observe the last decision made by the judge. We also assume that the wage w that the judge obtains per case (or the probability of an election) is high enough that for any sentence S and any case specific factors s_t, the judge prefers choosing sentence S to guarantee reelection in case of a next period election rather than taking an otherwise preferable decision that guarantees electoral defeat.

5.2 Results

The goal of this model is to understand how the planner will choose among the different tools to delegate the decision to the judge. The essence of the comparison is given in the following Proposition:

Proposition 1 *There exists $\tilde{V}$ and $\bar{V}$ such that:*

1. *if the variance of s_t , $V(s_t) \geq \tilde{V}$, rotation is the socially preferable instrument,*

2. *if $V(s_t) = 0$ and $V(\beta_d) \leq \bar{V}$, sentencing guidelines are socially preferable to rotation.*

The key tradeoff can be understood by describing the main instruments. All details can be found in the Appendix.

If the planner chooses **sentencing guidelines**, she sets the sentence optimally, without knowing s_t or β_d. Given this uncertainty and the fact these variables have zero mean, she sets $S_g = \theta$, and the realized welfare is

$$W_g = -(1+\alpha)V\left(s_t\right) - \alpha V\left(\beta_d\right),$$

welfare that decreases when the uncertainty on s_t and β_d increases. W_g is particularly sensitive to uncertainty on s_t, since better knowledge of local conditions allows the planner to better tailor the sentence to her preferences, but also better to minimize the cost of enforcement.

If the planner chooses **rotation**, she will choose the number of cases per rotation N sufficiently large for the judge to find it profitable to invest in learning the local norm. The benefit of rotation is that the judge takes into account the severity s_t and the norm β_d. The costs are twofold. First, the judge, will excessively adapt to the norm because she bears more cost than the planner from deviation. Second, with probability $1-p$ the judge is tough and has preferences that are not aligned with the planner.

The last instrument, **elections**, has very particular features. The equilibrium with elections is characterized by pooling on a single sentence that ignores the case specifics s_t (as in the comparison between elected and appointed civil servants, Maskin and Tirole, 2004). Indeed, a separating equilibrium cannot exist, since, in such equilibria, higher sentences would always be more likely to be chosen by a tough judge. Depending on whether the districts prefer a tough or a normal judge, there would always be a type willing to mimic the other. Thus, under an election system, sentences do not depend on s_t.

Proposition 1.1 then follows naturally. Since conditions s_t are ignored under both elections and sentencing guidelines, rotation dominates when the variance of s_t is large, i.e when there is a large variance in case details. Proposition 1.2 follows from the fact that when both the variance of s_t and β_d is small, the benefits of rotation disappear and there only remains the cost that the judge might be tough and have preferences not aligned with the planner.

In the model, as expressed in Proposition 1 a single instrument is socially preferred, and there is no room for complementarities between instruments. In practice, the judicial system in North Carolina uses a mix of all three instruments. However, the sentencing guidelines implemented in the state are actually fairly large intervals of acceptable sentences. In this sense, they could help achieve a better balance between the different dimensions of the trade-off. Indeed, under rotation, for relatively more socially acceptable crimes (high β_d), the judge adapts by giving lower sentences. This may be socially beneficial because enforcement costs can arise when the sentences are far from the social norms. However, there is also the risk that the judge adapts excessively to these local conditions. In this sense, loose sentencing guidelines, as those implemented in North Carolina where a relatively large interval of sentences is authorized for each crime, might be optimal to generate constrained adaptation.

6 Conclusion

In this paper, we have provided evidence on the existence of local sentencing norms and shown that judges outside their home district gradually converge to them. The results are robust to alternative definitions of the norm and different ways of defining the working sample. We also document factors affecting the speed of the convergence process.

When we discuss welfare implications in the model of Section 5, a key element is whether this convergence is considered to be socially desirable or not. In the model, some adaptation to local norms is desirable, because sentences that are in conflict with local customs are harder to enforce. This is consistent with the empirical finding that districts with more prevalent crime, which we interpret as being more socially acceptable, have lower sentencing norms.

References

Abrams, D.S., Bertrand, M. and Mullainathan, S.. 2012. "Do Judges Vary in Their Treatment of Race?" *Journal of Legal Studies*, 41(2): 347-383.

Abrams, D. and Fackler, R.. 2018. "To Plea or Not to Plea: Evidence from North Carolina." *Working Paper*, University of Pennsylvania.

Acemoglu, D. and Jackson, M.O.. 2017. "Social Norms and the Enforcement of Laws." *Journal of the European Economic Association*, 15(2): 245-295.

Blanes i Vidal, J. and Leaver, C.. 2011. "Are tenured judges insulated from political pressure?" *Journal of Public Economics*, 95(7-8): 570-586.

Bobbitt, W.H.. 1948. "Rotation of Superior Court Judges." *North Carolina Law Review*, 26: 335-349.

Bordalo, P., Gennaioli, N. and Shleifer, A.. 2015. "Salience Theory of Judicial Decisions." *Journal of Legal Studies*, 44(1): 7-33.

Besley, T. and Payne, A.A.. 2013. "Implementation of Anti-Discrimination Policy: Does Judicial Selection Matter." *American Law and Economics Review*, 15(1): 212-251.

Berdejo, C. and Chen, D.. 2017. "Electoral Cycles Among U.S. Courts of Appeals Judges." *Journal of Law and Economics*, 60(3): 479-496

Drago, F., Galbiati, R. and Vertova, P.. 2009. "The Deterrent Effect of Prison Evidence from a Natural Experiment." *Journal of Political Economy*, 117(2): 257-280.

Finkelstein, E., Gentzkow, M. and Williams, H.. 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Quarterly Journal of Economics*, 131(4): 1681-1726.

Cohen, A. and Yang, C.. 2018. "Judicial Politics and Sentencing Decisions." *American Economic Journal: Economic Policy*, Forthcoming.

Coviello, D. and Gagliarducci, S.. 2017. "Tenure in Office and Public Procurement." *American Economic Journal: Economic Policy*, 9(3): 59-105.

Coviello, D., Ichino, A. and Persico, N.. 2018. "Measuring the Gains from Labor Specialization: Theory and Evidence." *mimeo*, Northwestern University.

Dal Bo, E. and Rossi, M.. 2011. "Term Length and the Effort of Politicians." *Review of Economic Studies*, 78(4): 59-105.

Eisenstein, J., Flemming, R.B., Nardulli, P.F.. 1988. "The Contours of Justice: Communities and Their Courts." *Boston: Scott, Foresman.*

Gennaioli, N. and Shleifer, A.. 2008. "Judicial Fact Discretion." *Journal of Legal Studies*, 88(2):1-35.

Hay, J.R. and Shleifer, A.. 1998. "Private Enforcement of Public Laws: A Theory of Legal Reform." *American Economic Review*, 88(2): 398-403.

Ichino, A., Polo, M. and Rettore, E.. 2003. "Are Judges Biased by Labor Market Conditions?" *European Economic Review*, 47(5): 913-944.

Lim, C.. 2013. "Preferences and Incentives of Appointed and Elected Public Officials." *American Economic Review*, 103 (4):1360-1397.

Lim, C., Silveira, B. and Snyder, J.. 2016. "Do Judges' Characteristics Matter? Ethnicity, Gender, and Partisanship in Texas State Trial Courts." *American Law and Economics Review*, 18 (2): 302-357.

Maskin, E. and Tirole, J.. 2004. "The Politician and the Judge: Accountability in Government." *The American Economic Review*, 94 (4): 1034-1054

Mnookin, R.H. and Kornhauser, L.. 1979 "Bargaining in the Shadow of the Law: The Case of Divorce" *The Yale Law Journal*, 88(5): 950-997

Molitor, D.. 2018. "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration." *American Economic Journal: Economic Policy*, 10(1): 326-356.

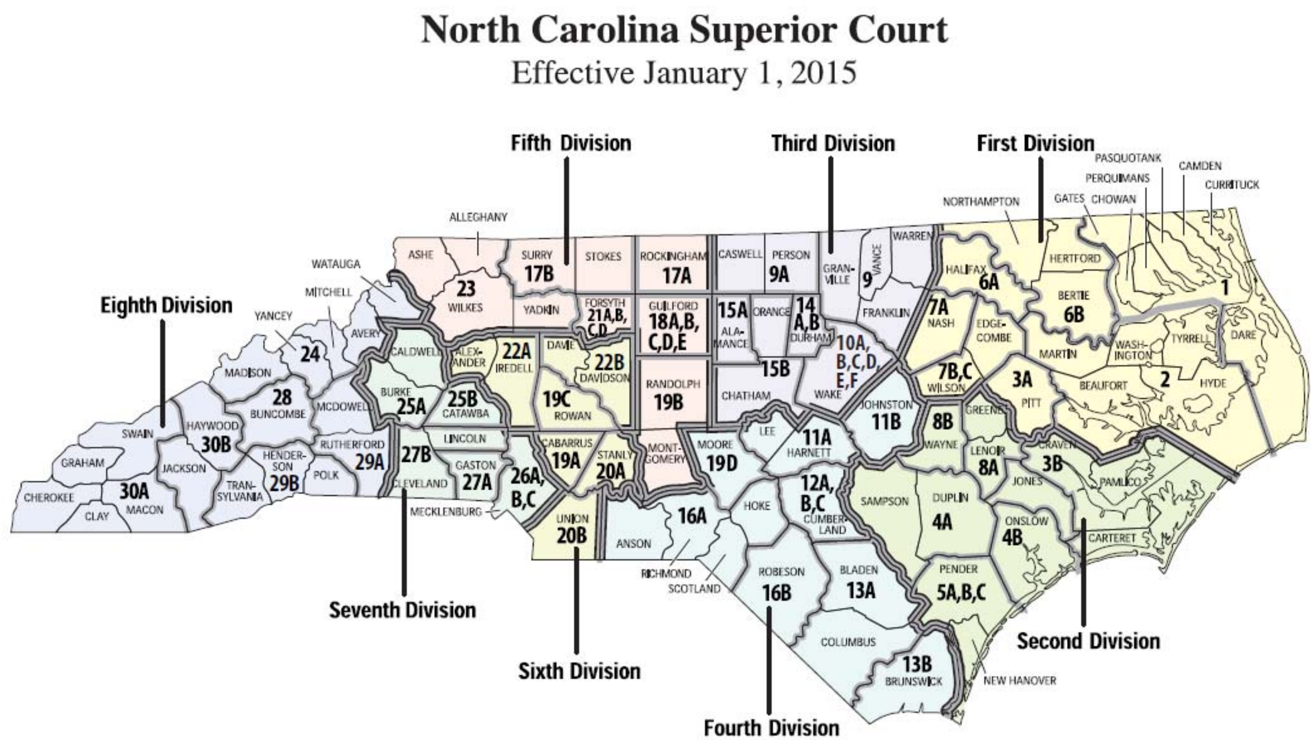
Schanzenbach, M. M. and Tiller, E.H.. 2007. "Strategic Judging under the U.S. Sentencing Guidelines: Positive Political Theory and Evidence." *Journal of Law, Economics and Organization*, 23(1): 24-56.

Ulmer J.T. and Johnson, B.. 2004. "Sentencing in Context: A Multilevel Analysis." *Criminology*, 42 (1):137-177.

Yang, C.S. 2014. "Have Interjudge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker." *NYU Law Review*, 89(4): 1268-1342.

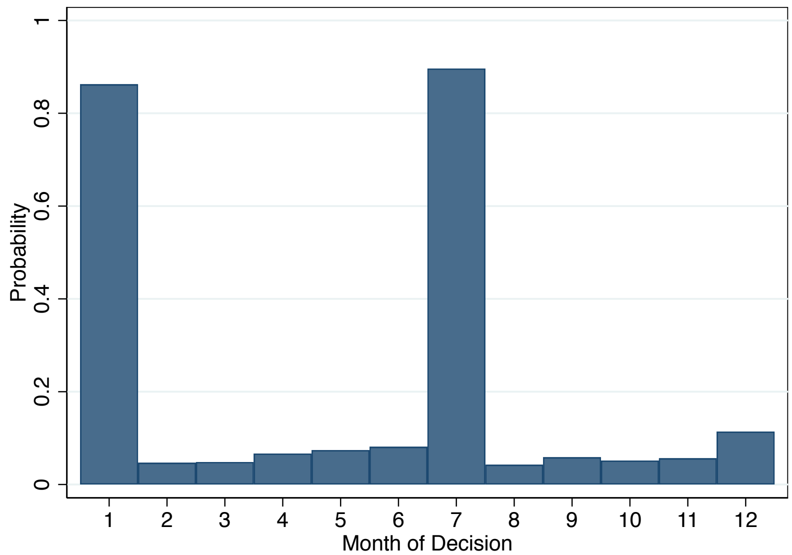
Appendix A

Figure 1: Judicial Map of North Carolina



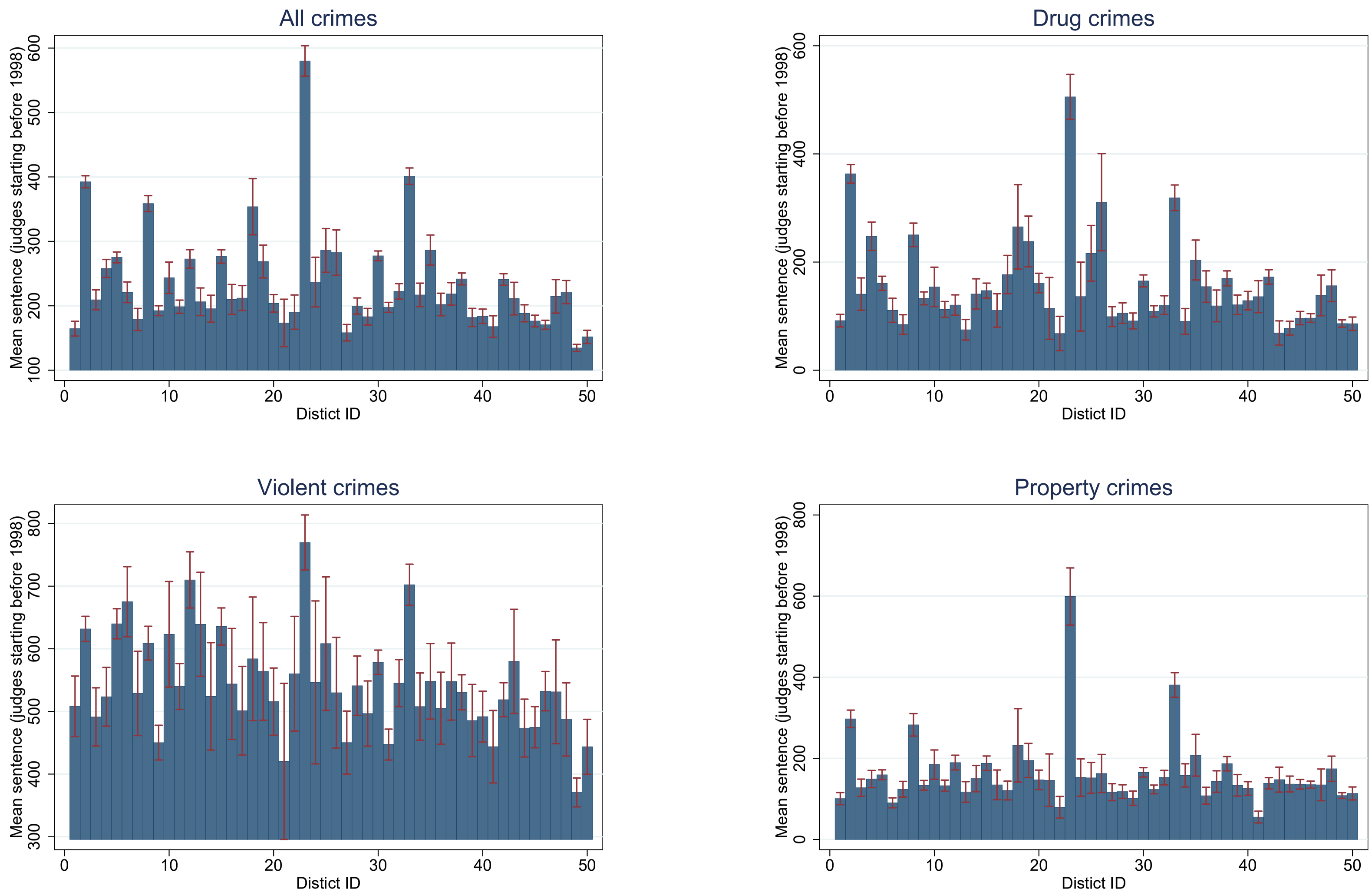
Note: Districts that have more than one letter associated with the district number (i.e., 10A, B, C, D) are divided into separate districts for electoral purposes. For administrative purposes, they are combined into a single district.

Figure 2:



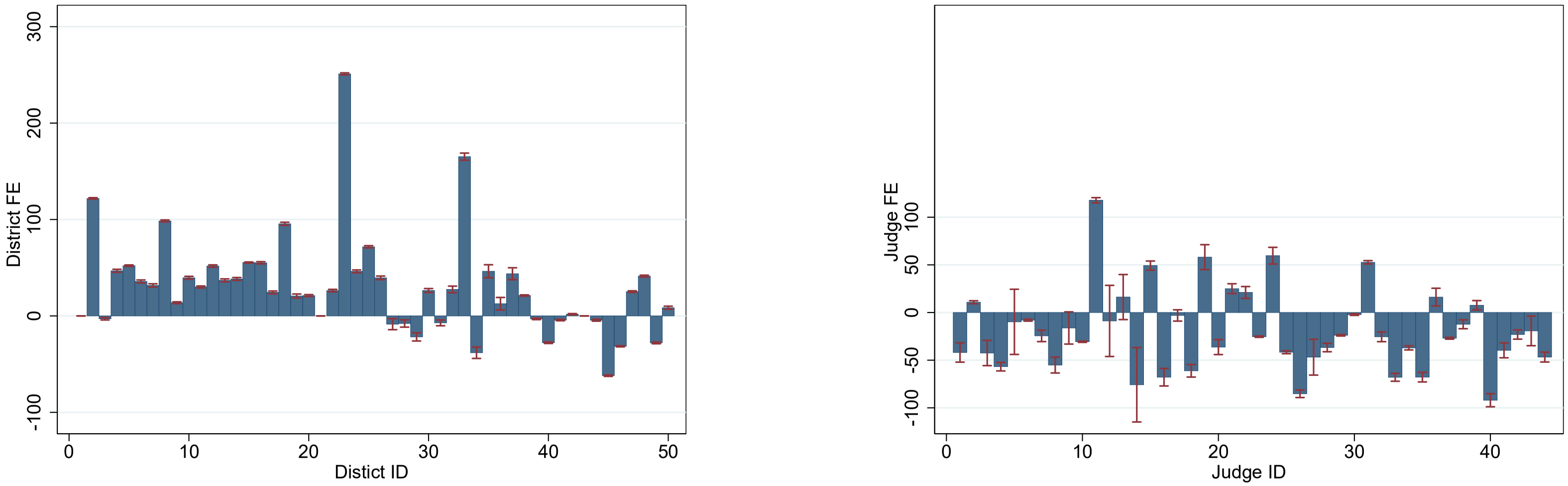
Note: Probability of moving per month, using data from the master schedule. A move is defined as a change in the district where the majority of decisions are made by a judge in a month.

Figure 3: Spatial Variation in Sentencing



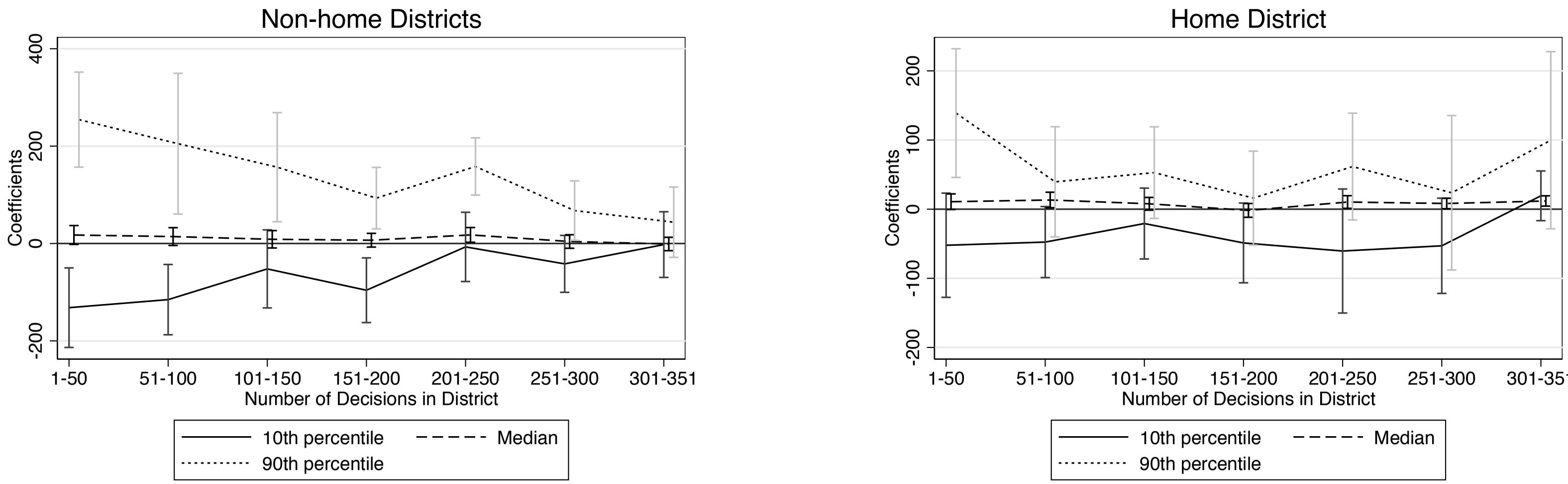
Note: Each figure shows the mean sentence by district among senior judges (those appointed before 1998) for: all crimes (top-left panel), drugs (top right), violent crimes (bottom left), and property crimes (bottom right)

Figure 4: Comparison of Spatial and Judicial fixed effects



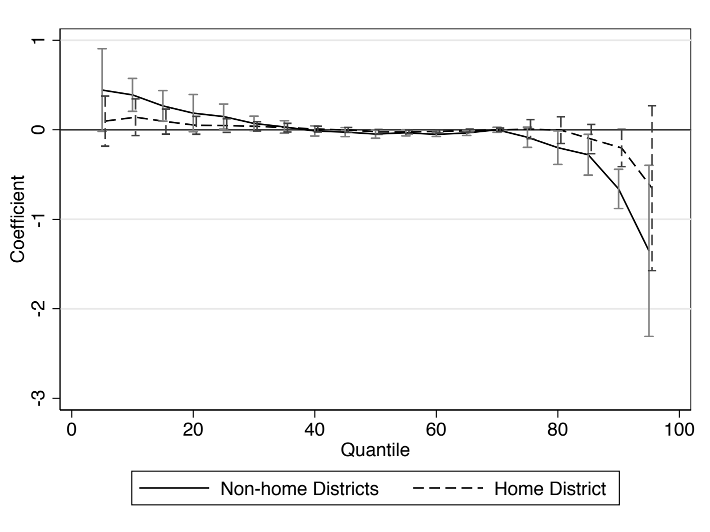
Note: The left panel displays the distribution of district fixed effects. The right panel displays judicial fixed effects from equation (1) estimated using only senior judges.

Figure 5: District Experience and Sentencing



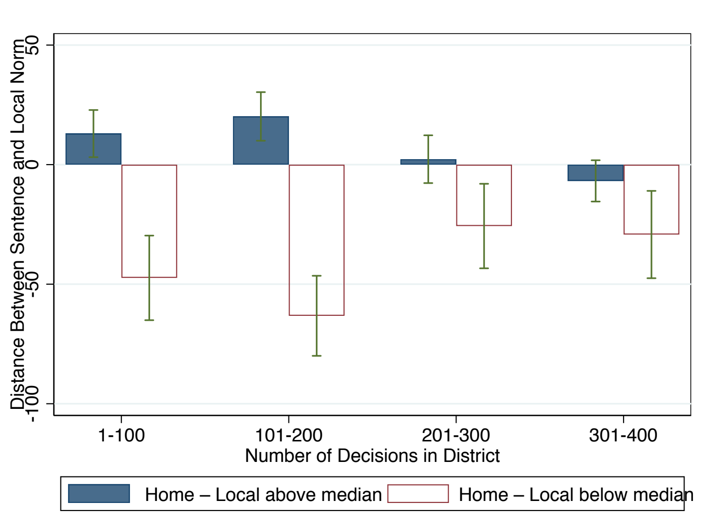
Note: Each panel presents results of three separate quantile regressions of the specification in Equation 2. In each case, the dependent variable is a measure of the relative severity of sentencing, and the independent variable is the number of cases in a district. The left panel presents results for judges outside their home district; the left for judges in their home district. Standard errors are clustered at the judge level and error bars indicate the 95% confidence interval.

Figure 6: District Experience and Sentencing, by quantile



Note: This figure presents the results of 19 different quantile regressions (from q5 on the left to q95 on the right). Each quantile regression measures the effect of the order of cases in a district on the quantile of the distance between sentences and the local sentencing norm. The dashed line presents estimates for judges in their home district and the solid line for judges in non-home districts. Standard errors are clustered at the judge level and error bars indicate the 95% confidence interval.

Figure 7: Impact of the Home District Local Norm



Note: Each bar presents the average distance between the assigned sentence and the local sentencing norm for judges not elected in the district. Judges with stricter (i.e. higher) home district local norms are represented by the blue bars on the left, and those with more lenient home districts are represented by the red bars on the right. Decisions are organized in groups of 100.

Table 1: Descriptive Statistics

Panel A: Case Characteristics

	Junior		Senior		All	
	mean	sd	mean	sd	diff	t
Defendant characteristics						
female	0.16	0.37	0.15	0.36	0.01	(7.73)
black	0.51	0.50	0.57	0.50	-0.06	(-32.24)
white	0.42	0.49	0.36	0.48	0.06	(35.19)
other race	0.08	0.05	0.11	0.08	-0.02	(-36.34)
minor	0.04	0.19	0.04	0.19	-0.00	(-0.52)
age	31.07	10.49	30.90	10.34	0.18	(4.85)
criminal history	1.68	1.15	1.65	1.13	0.03	(7.16)
first offense	0.46	0.50	0.45	0.50	0.01	(3.82)
Judge characteristics						
tenure	6.56	5.19	14.23	6.08	-7.67	(-378.90)
Judge in district of election	0.47	0.50	0.55	0.50	-0.08	(-47.79)
Outcomes						
Sentence (days)	501.85	1450.18	524.52	1569.44	-22.67	(-4.32)
active sentence (days)	360.45	1469.63	386.22	1586.64	-25.77	(-4.85)
Observations	198109		145667		343776	

Panel B: District Characteristics

	mean	sd	min	max
Judicial characteristics				
n. judge	11.5	3.97	6	25
n. district_attorney	21.66	14.96	7	96
n. attorney	61.08	39.45	13	192
prison overcrowding	.99	.13	.76	1.48
population characteristics				
share Democrat in presidential	.46	.07	.36	.59
unemployment rate	.04	.006	.03	.06
prop black	.22	.07	.07	.32
prop white	.69	.12	.39	.88
prop female	.52	.01	.47	.55
Observations	50	50	50	50

Table 2: Balancing tests

Panel A: Defendant Characteristics

	(1) First offense	(2) Female	(3) Black	(4) Age	(5) Minor	(6) Senior
Nb case in district	-0.0000756	-0.000000596	0.000192*	0.000348	0.000000504	0.00000221
	(0.0000555)	(0.0000344)	(0.000105)	(0.00110)	(0.0000227)	(0.00000383)
Nb case in district * Elected	0.0000200	0.0000185	0.000000574	0.000316	-0.00000575	-0.00000338
	(0.0000693)	(0.0000388)	(0.0000821)	(0.00133)	(0.0000247)	(0.00000518)
Elected	0.00294	0.00482	-0.0583*	0.474	-0.00922	0.00171
	(0.0194)	(0.00992)	(0.0347)	(0.360)	(0.00681)	(0.00114)
Nb case	-0.00000643	-0.00000560	-0.0000653***	0.000416***	-0.00000622***	0.000000415
	(0.00000777)	(0.00000543)	(0.0000211)	(0.000113)	(0.00000200)	(0.000000422)
Observations	33069	33069	33069	32469	32469	32469
Adjusted R^2	0.000	0.000	0.007	0.001	0.001	-0.000

Panel B: Crime Categories

	(1) Assault	(2) Burglary	(3) Drug	(4) Fraud	(5) Larceny	(6) Robbery
Nb case in district	-0.0000284	0.0000705	0.0000702	0.0000275	0.00000211	-0.0000424
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Nb case in district * Elected	0.0000194	-0.0000639	0.00000679	0.00000944	-0.000000228	0.0000328
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Elected	-0.00633	0.0110	-0.00975	-0.000931	-0.00483	-0.00738
	(0.005)	(0.014)	(0.020)	(0.009)	(0.009)	(0.010)
Nb case	0.00000605**	-0.0000179**	-0.00000548	-0.00000249	-0.00000384	0.00000907
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	33069	33069	33069	33069	33069	33069
Adjusted R^2	0.000	0.001	0.000	0.000	0.000	0.001

*Notes. Panel A: First offense is a dummy variable equaling one if it is the offender's first criminal case. Female and Black are dummies equaling one for defendants who are female and black, respectively. Minor and Senior are dummies equaling one if defendant is below 18 and above 64, respectively. Panel B: Each regression presents the effect on a dummy equaling one if the crime committed corresponds to the category mentioned in the header. In both panels: Case number in district is the order of the case in the district for the judge. Elected is a dummy equaling one if the judge is elected in the district. Case number is the order of the case over the judge's entire career. Standard errors are clustered at the judge level. t-Statistics are reported in parentheses. $p<0.1$, ** $p<0.05$, *** $p<0.01$*

Table 3: Impact of Case Order on Distance Between Sentence and Local Norm sentence

	(1) OLS	(2) q10	(3) q25	(4) q50	(5) q75	(6) q90	(7) $\lvert distance \rvert$
Case number in district	-0.064	0.390***	0.148*	-0.048*	-0.084	-0.661***	-0.236***
	(0.052)	(0.112)	(0.085)	(0.029)	(0.069)	(0.134)	(0.084)
Case number in district	0.049	-0.249*	-0.101*	0.024	0.090	0.457***	0.164***
* Elected	(0.056)	(0.146)	(0.060)	(0.028)	(0.056)	(0.127)	(0.055)
Elected	-6.974	17.573	26.304	10.652	-5.487	-63.161*	-25.419
	(13.527)	(34.429)	(17.756)	(7.666)	(5.560)	(33.392)	(15.149)
Case number	0.012	-0.084***	-0.020	0.010**	0.030	0.128**	0.046**
	(0.009)	(0.021)	(0.018)	(0.004)	(0.030)	(0.050)	(0.021)
Observations	33069	33069	33069	33069	33069	33069	33069

Notes. In columns 1 to 6, the outcome variable is the distance between the sentence and the average sentence, determined locally. Column 1 presents the result when using an OLS regression, while columns 2 to 6 present the results when using quantile regressions. The quantile is indicated in the header. In column 7, the outcome variable is the absolute value of the distance between the sentence and the average sentence, determined locally. Case number in district is the order of the case in the district for the judge. Elected is a dummy equal to one if the judge is elected in the district. Case number is the order of the case over the judge's entire career. Standard errors are clustered at the judge level and reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 4: Robustness of Convergence to Local Norm

	(1) Norm S elected	(2) Norm S+J>400	(3) 300	(4) 500	(5) S-D Controls	(6) Judge fe	(7) Active+inactive part
Case number in district	-0.356***	-0.241***	-0.155**	-0.250***	-0.244***	-0.248***	-0.211***
	(0.106)	(0.086)	(0.068)	(0.071)	(0.085)	(0.086)	(0.066)
Case number in district	0.261***	0.152***	0.152**	0.126***	0.170***	0.171***	0.135**
* Elected	(0.075)	(0.055)	(0.067)	(0.045)	(0.054)	(0.058)	(0.051)
Elected	-21.326	-22.234	-34.411**	-15.317	-24.770	-18.199	-20.432
	(23.524)	(14.899)	(16.970)	(17.562)	(14.789)	(20.593)	(13.614)
Case number	0.061**	0.047**	0.020*	0.067***	0.047**	0.049**	0.045***
	(0.029)	(0.021)	(0.012)	(0.017)	(0.021)	(0.021)	(0.016)
Observations	19361	33809	29372	33751	32469	33069	33069
Adjusted R^2	0.012	0.007	0.002	0.009	0.024	0.025	0.007

Notes. In every column, the outcome variable is the absolute value of the distance between the sentence and the average sentence, determined locally. In column 1, the average sentence is calculated among senior judges elected in the district. In column 2, the average sentence is calculated among all senior judges and junior judges after their 400th decision in the district. In column 3, we keep only the first 300 decisions made by a junior judge in a district. In column 4, we keep only the first 500 decisions made by a judge in a district. In column 5, we control for the following case characteristics: sex, race, age, and dummies for below 18 and above 64. In column 6, we control for judicial fixed effects. In column 7, the outcome variable is the distance between total (active plus inactive) sentence and the local sentencing norm, calculated using total sentences. Case number in district is the order of the case in the district for the judge. Elected is a dummy equaling one if the judge is elected in the district. Case number is the order of the case over the judge's entire career. Standard errors are clustered at the judge level. Standard errors are clustered at the judge level and reported in parentheses. *$p<0.1$, ** $p<0.05$, *** $p<0.01$.

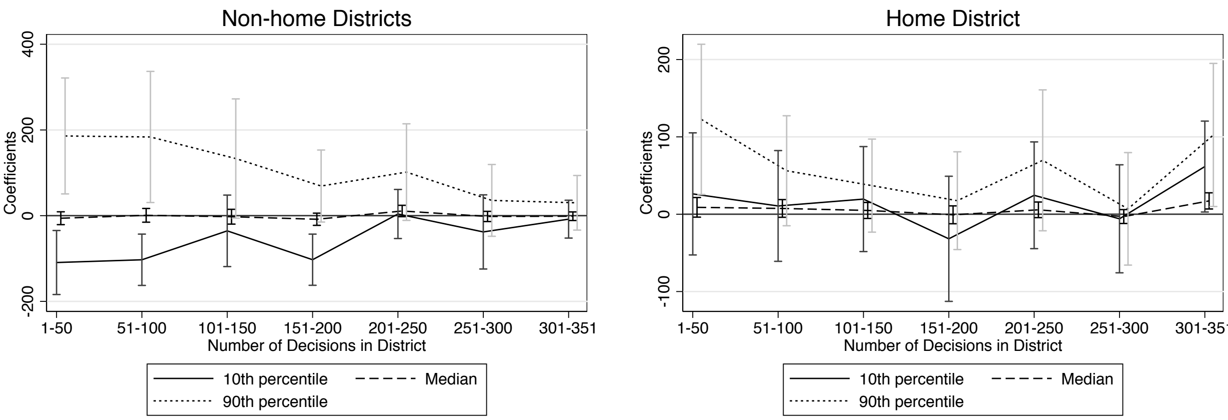
Table 5: Correlates of Local Norms

	Sentencing Norm				
	(1)	(2)	(3)	(4)	(5)
crime prevalence in d	-4.668***	-4.639***	-4.683***	-4.678***	-4.658***
	(1.572)	(1.618)	(1.575)	(1.578)	(1.583)
HI for attorneys			0.104	0.105	0.066
			(0.068)	(0.071)	(0.079)
HI for district attorneys			-0.023*	-0.021*	-0.007
			(0.012)	(0.011)	(0.011)
prison overcrowding				-43.353	-41.639
				(45.878)	(38.474)
majority democrat					29.949
					(25.126)
vote referendum 1996					1.114
					(3.593)
vote referendum 2004					-3.727*
					(2.023)
vote referendum 2010					-2.864
					(3.693)
vote referendum 2014					3.086
					(1.885)
Observations	948	948	948	948	948
Adjusted R^2	0.867	0.890	0.870	0.870	0.873

Note: All columns include crime fixed effects. We exclude crime categories that have fewer than 100 cases over the entire time period. Column (2) includes district fixed effects. Standard errors are clustered at the district level and reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

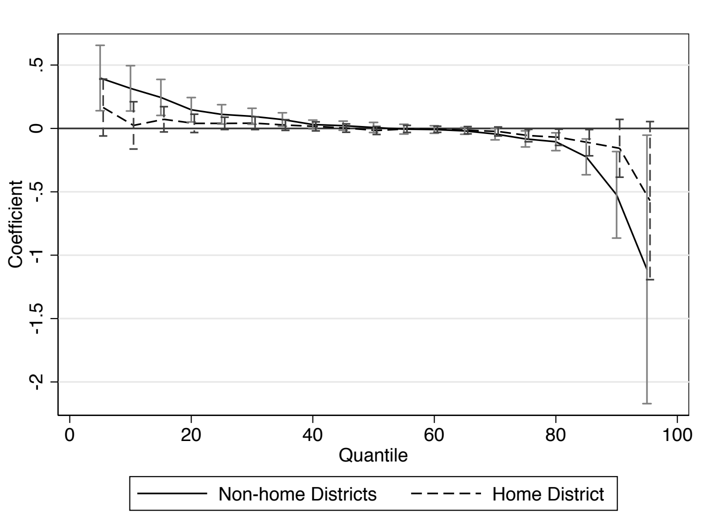
Appendix B: robustness using total sentence

Figure 8: Effect of decision order in the district on distance to the average sentence, using total sentences rather than active sentences



Notes. This figure reproduces Figure 5 for total sentences rather than active sentences

Figure 9: Effect of decision order in the district on distance for different quantiles, for total sentences



Notes. This figure reproduces Figure 6 for total sentences rather than active sentences

Table 6: Correlates of the sentencing norm, measured using total sentences

	Sentencing Norm				
	(1)	(2)	(3)	(4)	(5)
crime prevalence in d	-4.010***	-3.974***	-4.023***	-4.020***	-3.993***
	(1.416)	(1.454)	(1.418)	(1.419)	(1.425)
HI for attorneys			0.083	0.083	0.043
			(0.070)	(0.071)	(0.071)
HI for district attorneys			-0.023*	-0.022*	-0.003
			(0.012)	(0.012)	(0.011)
prison overcrowding				-19.576	-17.662
				(49.336)	(38.349)
majority democrat					37.367
					(24.640)
vote referendum 1996					1.769
					(3.663)
vote referendum 2004					-6.066***
					(2.053)
vote referendum 2010					-4.861
					(3.664)
vote referendum 2014					3.437*
					(1.804)
Observations	948	948	948	948	948
Adjusted R^2	0.870	0.899	0.873	0.872	0.878

Notes. This table reproduces Table 5 for total sentences rather than active sentences

Appendix C

We analyze the instruments in turn before proving the results of Proposition 1. We provide a lemma characterizing each instrument, with proofs provided later in the section.

Sentencing guidelines

The sentencing guideline is the sentence that maximizes the expected utility of the planner, uninformed about s_t and β_d, as expressed in the following lemma:

Lemma 1 *The sentencing guideline is $S_g = \theta$ and yields expected welfare of:*

$$W_g \quad = \quad -(1+\alpha)V\left[s_t\right] - \alpha V\left[\beta_d\right] \tag{3}$$

The expected welfare obtained under sentencing guidelines decreases when the uncertainty on s_t and β_d increase. Welfare is particularly sensitive to uncertainty on s_t, since better knowledge of local conditions allows the planner to better tailor the sentence to her preferences, but also to minimize the cost of enforcement.

Rotation policy

The judge has to decide whether to invest C to obtain information about the local norm β_d. This choice is determined by the difference in expected utility when informed about β_d and the expected utility without that information. The result is summarized below:

Lemma 2 *The planner rotates the judge every N^* periods, where N^* is defined as the lowest value of N such that:*

$$\frac{1-\delta^N}{1-\delta}\frac{\gamma^2}{1+\gamma}V(\beta_d) \geq C,$$

The planer under this rotation policy obtains a per period expected welfare of:

$$W_r = -\frac{\gamma^2+\alpha}{(1+\gamma)^2}V(\beta_d) - (1-p)(1+\alpha)\frac{1}{(1+\gamma)^2}\zeta^2 \tag{4}$$

The planner leaves the judge long enough in the district to ensure she acquires the information. If $V(\beta_d)$ increases, the judge rotates more frequently because information has a high value for the judge who no longer needs to be incentivized as much. The expected welfare of the planner under the rotation policy decreases with $V(\beta_d)$, as this increases the expected cost of enforcement, and decreases with the bias of the judge ζ.

Elections

We examine the final tool at the disposal of the planner, elections. Consistent with the empirical evidence we presented, showing that for judges in their home district there is no convergence to the sentencing norm, we assume that the judge, in her home district, knows β_d without having to invest C.

The citizen, in the case of an election in a soft district $\beta_d < 0$, wants to reelect a regular candidate and vote against a tough candidate. When focusing on the welfare-maximizing equilibrium, we obtain that:

Lemma 3 *If judicial election is the only tool at the planner's disposal, then all judges, regardless of their type or information on case specifics s_t, pool on a single sentence. If the judge pools on the socially preferred sentence, the expected welfare is given by:*

$$W_e = -\frac{\alpha}{(1+\alpha)}V(\beta_d) - (1+\alpha)V(s_t) \tag{5}$$

Lemma 3 shows that all judges, normal or tough, pool on a single sentence and ignore information on case specifics. Indeed, if there were at least two sentences chosen in equilibrium, the normal judge would always be more likely to choose the lowest one and would thus guarantee her own electoral defeat. In the welfare-maximizing equilibrium, the sentence on which they pool is the planner's preferred sentence $S^e = \theta + \frac{\alpha}{1+\alpha}\beta_d$.

Comparing instruments

Proposition 1 immediately follows from Lemmas 1-3. We restate the result:

Proposition 1 There exists $\tilde{V}$ and $\bar{V}$ such that:

1. if $V(s_t) \geq \tilde{V}$, rotation is the socially preferred instrument,
2. if $V(s_t) = 0$ and $V(\beta_d) \leq \bar{V}$, sentencing guidelines are socially preferred to rotation.

Lemma 2 shows that welfare under rotation optimal policy is independent of $V(s_t)$, while it is strictly decreasing both under sentencing guidelines and elections according to lemmas 1 and 3, so result 1 immediately follows, with a benchmark value $\tilde{V}$ of the variance of s_t, such that, if the variance is larger than that, rotation dominates. The key insight is that, in the pooling equilibrium, the judge ignores the local conditions s_t. The second result is based on how, according to Lemma 1, when $V(s_t) = 0$ and $V(\beta_d)$ converge to zero, welfare under sentencing guidelines goes to zero, while it is strictly negative for rotation due to the potential bias of the judge. Thus, there exists a critical value of the variance of β_d such that sentencing guidelines dominate.

Proof Lemma 1

For a given sentencing guideline S, the utility, given that s_t and β_d have zero means, is given by:

$$\begin{aligned}
E\left[u_p^d\right] &= -E\left[(S-(\theta+s_t))^2\right]-\alpha E\left[(S-(\theta+\beta_d+s_t))^2\right] \\
&= -S^2+2\theta S-\theta^2-E\left[s_t^2\right]-\alpha S^2+\alpha 2\theta S-\alpha E\left[s_t^2\right]-\alpha E\left[\beta_d^2\right]-\alpha\theta^2 \\
&= -(1+\alpha)S^2+2(1+\alpha)\theta S-(1+\alpha)\theta^2-(1+\alpha)V\left[s_t\right]-\alpha V\left[\beta_d\right]
\end{aligned}$$

The first order condition thus yields that the sentencing guideline is chosen as $S_g=\theta$. Thus, expected social welfare under sentencing guidelines is given by:

$$\begin{aligned}
W_g &= -V\left[s_t\right]-\alpha\left(V\left[s_t\right]+V\left[\beta_d\right]\right) \\
&= -(1+\alpha)V\left[s_t\right]-\alpha V\left[\beta_d\right]
\end{aligned}$$

Proof Lemma 2

We first determine the utility of the judge who invests in information and faces a case with severity s_t in a district with norm β_d. The sentence that solves the first order condition in this case is given by:

$$S_{r,i}=\theta+s_t+\frac{1}{1+\gamma}\zeta_j+\frac{\gamma}{1+\gamma}\beta_d$$

So the per period utility for the judge when she has information on β_d and s_t is:

$$\begin{aligned}
u_{r,i} &= -\left[\theta+s_t+\frac{1}{1+\gamma}\zeta_j+\frac{\gamma}{1+\gamma}\beta_d-(\theta+s_t+\zeta_j)\right]^2-\gamma\left[\theta+s_t+\frac{1}{1+\gamma}\zeta_j+\frac{\gamma}{1+\gamma}\beta_d-(\theta+s_t+\beta_d)\right] \\
&= -\left(\frac{\gamma}{1+\gamma}\right)^2(\beta_d-\zeta_j)^2-\gamma\left(\frac{1}{1+\gamma}\right)^2(\zeta_j-\beta_d)^2 \\
&= -\left(\frac{\gamma}{1+\gamma}\right)(\beta_d-\zeta_j)^2
\end{aligned}$$

So, the expected utility before obtaining the information is:

$$W_{r,i} = -\left(\frac{\gamma}{1+\gamma}\right)\left(V(\beta_d)+\zeta_j^2\right)$$

Now, we determine the expected welfare when no information on β_d is obtained. In this case, the sentence will be fixed at:

$$S_{r,u} = \theta + s_t + \frac{1}{1+\gamma}\zeta_j$$

For an expected welfare of:

$$W_{r,u} \quad = \quad -\left(\frac{\gamma}{1+\gamma}\right)\zeta_j^2 - \gamma V(\beta_d)$$

The per period welfare is naturally smaller when the judge is uninformed. There is thus a minimum number of periods N^r such that the judge will acquire information if and only if $N \geq N^r$, where N^r is defined is defined as the lowest value of N such that:

$$\frac{1-\delta^N}{1-\delta}\left(W_{r,i} - W_{r,u}\right) \geq C,$$

i.e

$$\frac{1-\delta^N}{1-\delta}\frac{\gamma^2}{1+\gamma}V(\beta_d) \geq C,$$

The planner will choose rotation after N periods and the welfare of the planner is given by:

$$W_r = -\frac{\gamma^2+\alpha}{(1+\gamma)^2}V(\beta_d) - (1-p)(1+\alpha)\frac{1}{(1+\gamma)^2}\zeta_j^2$$

Proof Lemma 3

Consider an equilibrium such that two level of sentences S_1 and $S_2 < S_1$ are awarded in equilibrium. There are benchmark value $\tilde{s}$ for the regular judge and $\bar{s}$ for the tough judge, with $\bar{s} < \tilde{s}$ such that the regular judge chooses S_1 if and only if $s_t \geq \tilde{s}$ and the tough judge chooses S_1 if and only if $s_t \geq \bar{s}$. Thus, since the regular type prefers lower sentences on average, when the voter sees sentence S_2 chosen, she increases her posterior belief that the incumbent is not tough and does not reelect her. Under our assumption that the wage is high enough that the judge wants to win the election at all costs, sentence S_2 would never be chosen in equilibrium.

Therefore, the unique equilibrium is a pooling equilibrium. In the welfare-maximizing equilibrium, the judges pool on the preferred sentence of the planner $S^e = \theta + \frac{\alpha}{1+\alpha}\beta_d$. The welfare is then given by:

$$W_e = -\frac{\alpha}{(1+\alpha)}V(\beta_d) - (1+\alpha)V(s_t)$$

Proof Proposition 1

1. According to lemmas 1 and 3, expected welfare under rotation and elections (even in the case of the pooling equilibrium that maximizes welfare) is decreasing in $V(s_t)$, while lemma 2 shows that welfare under rotation is independent of $V(s_t)$. Result (1) directly follows.

2. According to expression (3), when $V(s_t) = 0$ and $V(\beta_d)$ converges to 0, expected welfare under sentencing guidelines goes to zero. On the contrary, according to expression (4), expected welfare under rotation converges to $W_r = -(1-p)(1+\alpha)\frac{1}{(1+\gamma)^2}\zeta^2 < 0$. Thus there is a range of values of $V(\beta_d)$ such that sentencing guidelines dominate rotation.

Appendix D- Web Appendix

7 Variable Description and Data Set Construction

7.1 Case Definition: Charge and Sentence

The first step of our analysis is a case definition. Since a criminal case is often comprised of multiple charges for a single defendant, and our focus is on overall sentencing for a case, we build on the procedure described in Abrams and Fackler (2018) in identifying cases with the same criminal and disposition date, and we define a case as a unique person-date of disposition. Treating a multiple charges case as a single unit implies that we must decide which charge to keep. We proceed as follows: we first define the lead charge of an incident as the charge with the highest associated sentence length. Our main sentencing variable is then defined as the minimum sentence determined by the judge for the lead charge. As we have specified in the main text, when a defendant is found guilty to a felony, North Carolina imposes a sentencing range. If the judge determines the sentence should be active, the defendant is required to serve the full minimum of the range, and may serve less than the maximum with good behavior. The final active sentence is the main variable used in our analysis. It is worth noting that, in order to deal with the outliers in sentences, in our analysis we winsorize this variable at the 5 percent level.

7.2 Judge identity

The second step to conduct our analysis is to identify the judge dealing with each case. North Carolina sentencing data reports a judge acronym (with two or three letters) for each case. In order to identify a specific judge based on these acronyms, we use the Master Schedules recording in which district and division a judge is in a given week. Using this information, we construct one, two, and three letter acronyms for each judge in the schedule and match this with our case data. Using disposition date data and the acronyms, we are able to match the 84 percent of judges in the master schedule to cases in the sentencing data. We only keep these observations in our working sample. For judges elected or nominated after the 1998, we observe the whole history of decisions.

7.3 District level demographic variables

We collect various demographic and other district level variables that we use in different steps of the analysis. District level demographic characteristics are constructed starting from the

US Bureau of Census data. We use variables for the year 2010 (the most recent census fully available). These variables (listed in descriptive statistics table of the paper) are collected at the county level and are then aggregated at the district level since each district usually includes more than one county.

7.4 District level administrative and political variables

Prison population data are collected from the National Census of Jails and the Annual Survey of Jails and are used to construct a crowding metric as the ratio between the value of the total population of inmates at that county's jail facilities at the survey date and the rated capacity of the jail, which measures the maximum number of beds (and therefore overnight inmates) that could fit into the facility on the date the survey was taken.

Finally, we collect data on referenda about justice. We identified four referenda that took place in 1996, 2004, 2010, and 2014. The 1996 referendum asked voters about the expansion of alternative punishments to be used on convicted criminals, such as probation and community service. The 2004 referendum aims at clarifying and defining several areas of jurisdiction of the courts, and changed the term of office of magistrates to provide for an initial term of 2 years and subsequent terms of 4 years. The 2010 is intended to prohibit convicted felons for running as sheriffs in the state and finally the 2014 introduces the possibility for felons to waive a trial by jury. We collect data about county level votes in these referenda from North Carolina's Board of Elections and then aggregate them at the district level to compute the percentage of votes in favor or against the main question asked in the referendum.

Figure 10: Sentencing guidelines

Figure A

*** Effective for Offenses Committed on or after 10/1/13 ***

FELONY PUNISHMENT CHART

PRIOR RECORD LEVEL

OFFENSE CLASS	I 0-1 Pt	II 2-5 Pts	III 6-9 Pts	IV 10-13 Pts	V 14-17 Pts	VI 18+ Pts	
A	**Death or Life Without Parole Defendant Under 18 at Time of Offense: Life With or Without Parole**						
B1	**A**	**A**	**A**	**A**	**A**	**A**	**DISPOSITION**
	240 - 300	*276 - 345*	*317 -397*	*365 - 456*	*Life Without Parole*	*Life Without Parole*	*Aggravated Range*
	192 - 240	**221 - 276**	**254 - 317**	**292 - 365**	**336 - 420**	**386 - 483**	**PRESUMPTIVE RANGE**
	144 - 192	*166 - 221*	*190 - 254*	*219 - 292*	*252 - 336*	*290 - 386*	*Mitigated Range*
B2	**A**	**A**	**A**	**A**	**A**	**A**	
	157 - 196	*180 - 225*	*207 - 258*	*238 - 297*	*273 - 342*	*314 - 393*	
	125 - 157	**144 - 180**	**165 - 207**	**190 - 238**	**219 - 273**	**251 - 314**	
	94 - 125	*108 - 144*	*124 - 165*	*143 - 190*	*164 - 219*	*189 - 251*	
C	**A**	**A**	**A**	**A**	**A**	**A**	
	73 – 92	*83 - 104*	*96 - 120*	*110 - 138*	*127 - 159*	*146 - 182*	
	58 - 73	**67 - 83**	**77 - 96**	**88 - 110**	**101 - 127**	**117 - 146**	
	44 - 58	*50 - 67*	*58 - 77*	*66 - 88*	*76 - 101*	*87 - 117*	
D	**A**	**A**	**A**	**A**	**A**	**A**	
	64 - 80	*73 - 92*	*84 - 105*	*97 - 121*	*111 - 139*	*128 - 160*	
	51 - 64	**59 - 73**	**67 - 84**	**78 - 97**	**89 - 111**	**103 - 128**	
	38 - 51	*44 - 59*	*51 - 67*	*58 - 78*	*67 - 89*	*77 - 103*	
E	**I/A**	**I/A**	**A**	**A**	**A**	**A**	
	25 - 31	*29 - 36*	*33 - 41*	*38 - 48*	*44 - 55*	*50 - 63*	
	20 - 25	**23 - 29**	**26 - 33**	**30 - 38**	**35 - 44**	**40 - 50**	
	15 - 20	*17 - 23*	*20 - 26*	*23 - 30*	*26 - 35*	*30 - 40*	
F	**I/A**	**I/A**	**I/A**	**A**	**A**	**A**	
	16 - 20	*19 - 23*	*21 - 27*	*25 - 31*	*28 - 36*	*33 - 41*	
	13 - 16	**15 - 19**	**17 - 21**	**20 - 25**	**23 - 28**	**26 - 33**	
	10 - 13	*11 - 15*	*13 - 17*	*15 - 20*	*17 - 23*	*20 - 26*	
G	**I/A**	**I/A**	**I/A**	**I/A**	**A**	**A**	
	13 - 16	*14 - 18*	*17 - 21*	*19 - 24*	*22 - 27*	*25 - 31*	
	10 - 13	**12 - 14**	**13 - 17**	**15 - 19**	**17 - 22**	**20 - 25**	
	8 - 10	*9 - 12*	*10 - 13*	*11 - 15*	*13 - 17*	*15 - 20*	
H	**C/I/A**	**I/A**	**I/A**	**I/A**	**I/A**	**A**	
	6 - 8	*8 - 10*	*10 - 12*	*11 - 14*	*15 - 19*	*20 - 25*	
	5 - 6	**6 - 8**	**8 - 10**	**9 - 11**	**12 - 15**	**16 - 20**	
	4 - 5	*4 - 6*	*6 - 8*	*7 - 9*	*9 - 12*	*12 - 16*	
I	**C**	**C/I**	**I**	**I/A**	**I/A**	**I/A**	
	6 - 8	*6 - 8*	*6 - 8*	*8 - 10*	*9 - 11*	*10 - 12*	
	4 - 6	**4 - 6**	**5 - 6**	**6 - 8**	**7 - 9**	**8 - 10**	
	3 - 4	*3 - 4*	*4 - 5*	*4 - 6*	*5 - 7*	*6 - 8*	

A – Active Punishment **I – Intermediate Punishment** **C – Community Punishment**
Numbers shown are in months and represent the range of minimum sentences

Revised: 09-09-13

Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex). Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR). Il est financé dans le cadre des investissements d'avenir.

(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)

www.sciencespo.fr/liepp

A propos de la publication

Procédure de soumission :
Rédigé par un ou plusieurs chercheurs sur un projet en cours, le *Working paper* vise à susciter la discussion scientifique et à faire progresser la connaissance sur le sujet étudié. Il est destiné à être publié dans des revues à comité de lecture (peer review) et à ce titre répond aux exigences académiques. Les textes proposés peuvent être en français ou en anglais. En début de texte doivent figurer : les auteurs et leur affiliation institutionnelle, un résumé et des mots clefs.

Le manuscrit sera adressé à : liepp@sciencespo.fr

Directeur de publication :
Bruno Palier

Comité de rédaction :
Andreana Khristova, Carolina Alban Paredes

Sciences Po - LIEPP
27 rue Saint Guillaume
75007 Paris - France
+33(0)1.45.49.83.61
liepp@sciencespo.fr